Obtención de la media aritmética datos no tabulados. La media aritmética se define como la suma de los valores de la variable dividida entre el número de datos, lo cual se expresa mediante la siguiente ecuación:

Remarca: En estadística descriptiva para su cálculo por el momento denotaremos a la media aritmética lo denotaremos por μ

$$\mu = \sum_{i=1}^{n} \frac{x_i}{n} = \frac{x_1 + x_2 + \cdots + x_n}{n}, \quad x_i = datos$$

Ejemplo: Si tenemos los siguientes datos: 2, 4, 5, 8, 10, 10
Determine la media aritmética

$$\mu = 6.5$$

Sol: Para determinar la media apliquemos la siguiente expresión: $\mu = \frac{\sum X_i}{n}$

$$\mu = \frac{\sum X_i}{n} = \frac{X_1+X_2+X_3+X_4+X_5+X_6}{6} = \frac{2+4+5+8+10+10}{6} = 6.5$$

Obtención de la media aritmética datos tabulados. La media aritmética cuando los datos están tabulados, se obtiene aplicando una de las siguientes expresiones:

$$\mu = \sum_{i=1}^{m} \frac{x_i * n_i}{n} = \frac{x_1 * n_1 + x_2 * n_2 + \cdots + x_m * n_m}{n}, \quad m = numero\ de\ filas$$

$$\mu = \sum_{i=1}^{m} x_i * h_i$$

Ejemplo.-

Cuadro #
Tabla de distribución de frecuencias

[L_{i-1} - L_i >	**Xi**	**n_i**	**h_i**	**h_i (%)**	**N_i**	**H_i (%)**
[1.3 - 1.84)	1.57	8	0.4	40%	8	40%
[1.84 - 2.38)	2.11	3	0.15	15%	11	55%
[2.38 - 2.92)	2.65	6	0.3	30%	17	85%
[2.92 -3.44]	3.18	3	0.15	15%	20	100%
TOTAL		n=20	1	100%		

$$\mu = \sum_{i=1}^{4} \frac{x_i * n_i}{n} = \frac{x_1 * n_1 + x_2 * n_2 + x_3 * n_3 + x_4 * n_4}{n}$$

$$= \frac{1.57 * 8 + 2.11 * 3 + 2.65 * 6 + 3.18 * 3}{20} = 2.216$$

Remarca: Para el cálculo de la media aritmética se usará con tres dígitos después del punto decimal

Cabe resaltar que utilizando las expresiones precedentes cuando tenemos la distribución tipo III, obtenemos una aproximación del verdadero valor de la media aritmética.

Propiedades de la media aritmética. La media aritmética como medida de posición goza de las siguientes propiedades:

- ✓ La media aritmética de una constante es la misma constante.

$$\mu = \mu_k = M(k) = k, \quad k = constante$$

Ejemplo: sean los datos: 3,3,3,3,3,3 $\mu = 3$

- ✓ La media aritmética de una variable multiplicada por una constante, es igual a la constante por la media de la variable.

$$\mu_{kx} = M(kx) = k * M(x)$$

Ejemplo: Sean los datos 5,10, 25

Determine la media aritmética

$$\mu_{kx} = 13.333$$

Forma solución 2:

$5 = 5 * 1$

$10 = 5 * 2$

$25 = 5 * 5$

$\mu_{kx} = M(kx) = k * M(x) = 5 * 2.667 = 13.333$

$M(x) = 2.67$

- ✓ La media aritmética más/menos una constante, es igual a la media aritmética de la variable más/menos la constante.

$$\mu_{k \pm x} = M(k \pm x) = k \pm M(x)$$

- ✓ Si una variable X, es particionable mediante 'r' variables, es decir:

$$x = x_1 \pm x_2 \pm x_3 \pm \cdots \pm x_r \qquad \forall\, x, Variables\ Independientes$$

entonces cumple:

$$M(x) = M(x_1 \pm x_2 \pm x_3 \pm \cdots \pm x_r) = M(x_1) \pm M(x_2) \pm \cdots \pm M(x_r)$$

Ventajas de la media aritmética. La media aritmética tiene las siguientes ventajas:

- ✓ Es un concepto intuitivo y conocido para la mayoría de las personas.
- ✓ Es una medida única, puesto que cada conjunto de datos solo tiene una media aritmética.
- ✓ Para su cálculo se toman en cuenta todas las observaciones del conjunto de datos.

Desventajas de la Media Aritmética. El uso de la media aritmética presenta las siguientes desventajas:

- ✓ Pueden estar afectada por valores extremos que no son representativos del resto de las observaciones.
- ✓ El cálculo el muchos casos puede resultar tedioso.

La Moda. La moda es una medida de posición que corresponde al valor de la variable o modalidad del atributo determinada por las frecuencias (absolutas o relativas) con mayor valor, la moda se denota frecuentemente por: Mo, Mo(x), etc

En el caso que una distribución de frecuencias tuviera una sola moda se denomina unimodal, si tiene dos modas bimodal, y si tiene más de tres modas multimodal.

Datos no tabulados

Ejemplo: si tenemos los datos: 2,2,2,4,5,6. Determine la moda

La moda es 2 (unimodal)

Ejemplo: si tenemos los datos: 2,2,4,4,5,5,6. Determine la moda

La moda es $Mo = 2, Mo = 4, Mo = 5$ (trimodal)

Datos tabulados

Calculo de la Moda. Para la obtención de la moda se consideran los dos siguientes casos:

Primer proceso. Se aplica en el caso de estar presentes frente a una distribución de tipo I o una distribución de frecuencias de un atributo, siendo la moda el valor de la variable o de la modalidad del atributo correspondiente a la frecuencia con mayor valor.

Ejemplo: Sea tabla, determine la moda

Cuadro # 3
Tabla de distribución de frecuencias
Peso de estudiantes (X)

X_i	n_i	hi	h_i %	N_i	H_i%
55	4	0.1143	11.43%	4	11.43%
57	4	0.1143	11.43%	8	22.86%
58	12	0.3428	34.28%	20	57.14%
63	8	0.2285	22.85%	28	80%
67	7	0.2	20%	35	100%
Total	**35**	**1**	**100%**		

Mo = 58

Segundo Proceso. Se aplica cuando los datos se han agrupado en una distribución tipo II, y se utiliza la siguiente expresión matemática.

$$Mo = L_{j-1} + c_j\left(\frac{n_j - n_{j-1}}{(n_j - n_{j-1}) + (n_j - n_{j+1})}\right), \qquad n_j = frecuencia\ absoluta\ maxima$$

Donde: n_j = Frecuencia absoluta con mayor valor

c_j = Longitud del intervalo de clase

L_{j-1} = Límite inferior correspondiente al intervalo con la frecuencia absoluta con mayor valor.

Ejemplo: Determine la moda

Ejemplo: determinar la moda

[L_{i-1} - L_i >	Xi	n_i	h_i (%)	N_i	H_i (%)
7 a 9.4	8.2	8	32	8	32
9.4 a 11.8	10.6	5	20	13	52
11.8 a 14.2	13	4	16	17	68
14.2 a 16.6	15.4	6	24	23	92
16.6 a 19	17.8	2	8	25	100
TOTAL		25	100		

$$Mo = L_{j-1} + c_j\left(\frac{n_j - n_{j-1}}{(n_j - n_{j-1}) + (n_j - n_{j+1})}\right)$$

En nuestro caso es:

$$Mo = 7 + 2.4\left(\frac{8 - 0}{(8 - 0) + (8 - 5)}\right) = 8.745$$

Remarca: para el cálculo usaremos con tres dígitos después del punto decimal

Ventajas de la moda. Las principales ventajas de la moda son:

- La moda se puede utilizar como una medida de posición para datos cualitativos como cuantitativos.
- La moda no está afectada por los valores extremos.

Desventajas de la moda. Las principales desventajas de la moda son:

- El hecho de poder tener dos o más modas puede ser inconveniente al momento de interpretar y comparar resultados

La mediana. La mediana se define como el valor de la variable tal que, si se ordenan los valores en forma creciente o decreciente, divide en dos partes iguales a la distribución respecto a la cantidad. La mediana frecuentemente se denota como: Me, Me(x), etc.

Calculo de la mediana. En el cálculo de la mediana es posible distinguir tres procesos.

Primer Proceso (no tabulados). Se utiliza para el caso de que los valores de la variable no se encuentren agrupados o tabulados. En este caso se aplica el concepto de la mediana para determinar su valor, considerando los siguientes aspectos.

- Se ordenan los valores no agrupados en forma ascendente o descendente.
- Si el número de datos es impar el valor de la mediana corresponde al valor de la variable que ocupa la posición central.

$$Me = X_{\left(\frac{n+1}{2}\right)}$$

- Si el número de datos es par el valor de la mediana es el promedio de los valores centrales de la variable.

$$Me = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2}$$

Ejemplo: Sean los datos

2, 3, 1, 7

Solución

Se ordena 1, 2, 3, 7 son una cantidad par $n = 4$

$$Me = \frac{X_{\frac{n}{2}} + X_{\frac{n}{2}+1}}{2} = \frac{X_{\frac{4}{2}} + X_{\frac{4}{2}+1}}{2} = \frac{X_2 + X_3}{2} = \frac{2+3}{2} = 2.5$$

Ejemplo: Sean los datos

2, 3, 1, 7, 11

Solución

Se ordena 1, 2, 3, 7, 11 son una cantidad impar n=5

$$Me = X_{\left(\frac{n+1}{2}\right)} = X_{\left(\frac{5+1}{2}\right)} = X_3 = 3$$

Datos tabulados

Segundo Proceso. Se aplica para los datos agrupados en la distribución tipo I, y para su cálculo se siguen los siguientes pasos:

- Se determina la cantidad media del universo, es decir: n/2
- Se ubica el valor n/2, entre los valores de la frecuencia absoluta acumulada, tal que el límite superior sea mayor o igual a n/2

  En tal caso se presentan dos situaciones:
  - Si $N_{i-1} < \frac{n}{2} < N_i$ entonces $N_i > \frac{n}{2}$, entonces el valor de la mediana es igual al valor correspondiente x_i, es decir:

$$Me = X_{i+1}$$

  - Si $N_i = \frac{n}{2}$, entonces el valor de la mediana es el promedio de los valores centrales de la variable, es decir:

$$Me = \frac{X_i + X_{i+1}}{2}$$

Ejemplo: determine la mediana

Si tenemos el siguiente cuadro de distribución de frecuencias, determinar la mediana

X_i	n_i	N_i
20	12	12
30	10	22
50	20	42 $\leftarrow \frac{n}{2} = 46$
60	18	60
70	15	75
80	10	85
90	7	92
Total	92	

Como $N_i > \frac{n}{2}$, en nuestro caso 60> 46, entonces la mediana está dada por: $Me = x_{i+1}$ que en nuestro caso es Me = 60.

Tercer Proceso Se utiliza cuando se manejan distribuciones tipo II y se siguen los pasos que a continuación se detallan:

- Se determina la cantidad media del universo, es decir. n/2.
- Se ubica el valor n/2 entre dos valores consecutivos de la frecuencia acumulada absoluta, tal que el límite superior sea mayor o igual a n/2.

  En tal caso se pueden presentar dos posibles casos:
  - Si $N_{i-1} < \frac{n}{2} < N_i$ entonces $N_i > \frac{n}{2}$ entonces el valor de la mediana se determina mediante la siguiente expresión.

$$Me = L_{i-1} + c_i\left(\frac{\frac{n}{2} - N_{i-1}}{N_i - N_{i-1}}\right)$$

- Si $N_i = \frac{n}{2}$ entonces el valor de la mediana corresponde al límite inferior del intervalo al cual pertenece el valor de N_i, es decir:

$$Me = L_{i-1}$$

Ejemplo: Determine la mediana

8. si tenemos la siguiente tabla determinar la mediana

[L_{i-1} - L_i >	**Xi**	**n_i**	**h_i (%)**	**N_i**	**H_i (%)**
7 a 9.4	8.2	8	32	8	32
9.4 a 11.8	10.6	5	20	13	52
11.8 a 14.2	13	4	16	17	68
14.2 a 16.6	15.4	6	24	23	92
16.6 a 19	17.8	2	8	25	100
TOTAL		25	100		

Sol: para determinar la mediana, seguimos los siguientes pasos

Paso 1. Determinar $\frac{n}{2}$, que en nuestro caso es $\frac{n}{2} = \frac{25}{2} = 12.5$

Paso 2 Ubicamos $\frac{n}{2}$, en los N_i

Paso 3. Trabajamos con la fila de debajo de donde está ubicado $\frac{n}{2}$, es decir, donde está encerrado y aplicamos la siguiente expresión:

$$Me = L_{j-1} + C_j\left(\frac{\frac{n}{2} - N_{j-1}}{N_j - N_{j-1}}\right)$$

$$Me = 9.4 + 2.4\frac{(12.5 - 8)}{(13 - 8)} = 11.560$$

Ventajas de la mediana. Las ventajas de utilizar la mediana son:

- La mediana es fácil de entender y puede ser calculada a cualquier clase de datos.
- La mediana es afectada por el número de observaciones y no por la presencia de los valores extremos.

4.3.3. Desventajas de la mediana. Las desventajas de la mediana son:

- Se deben organizar u ordenar los datos antes de cualquier tipo de cálculo. Esto es poco práctico cuando se tiene muchos datos.
- La mediana no es adecuada para efectuar manipulaciones algebraicas posteriores.

Análisis entre las medidas de posición media, mediana y moda

Datos Atípicos. - Se denomina un dato atípico a aquel valor que es extrañamente mucho más grande o mucho más pequeño que los datos tienen por lo general

Ejemplo: sean los datos

1, 500, 502, 503

Dato atípico: 1

Ejemplo: Sean los datos

3,4,5,6,7, 100

Dato atípico: 100

La media Aritmética. - Algunas características que debe cumplir los datos para considerar a la media aritmética como medida de posición representativa de los datos

1.- La mayor parte de los datos son próximos al valor central

2.- Por lo general no se tienen datos atípicos

3.- Los datos tienen un comportamiento similar a un comportamiento simétrico

4.- El valor de la media debe ser similar a varios datos

La mediana. - Algunas características que debe cumplir los datos para considerar a la mediana como medida de posición representativa de los datos

1.- El valor de la mediana debe ser similar o próximo a varios datos

2.- Si el valor de la mediana no es afectado por los datos atípicos, por lo cual el valor de la mediana no es afectado por los datos atípicos que puedan existir.

La moda. - Algunas características que debe cumplir los datos para considerar a la moda como medida de posición representativa de los datos

1.- El valor de la moda representa un porcentaje elévalo o igual a otros valores

2.- Por lo general tenemos un solo valor de la moda

Ejemplo: Sean los datos

2, 2.1, 3, 4, 4.3, 4.4, 5, 6, 7

¿Qué medida de posición mejor representa a los datos, si tenemos como alternativa la media, la mediana y moda?

Solución

$\mu = 4.2$ Esto debido a que no existe datos atípicos

$Me = 4.3$

$Mo = 2, Mo = 2.1, \ldots. \quad , Mo = 7$

Ejemplo: Sean los datos

3, 3, 9, 14, 16, 97

¿Qué medida de posición mejor representa a los datos, si tenemos como alternativa la media, la mediana y moda?

$\mu = 23.667$

$Me = 11.5$ porque existe datos atípicos

$Mo = 3$

Ejemplo: Sean los datos

2, 3, 3, 4, 4, 4, 4, 4

¿Qué medida de posición mejor representa a los datos, si tenemos como alternativa la media, la mediana y moda?

Media =3.5

Mediana= 4

Moda= 4